

Panoptical View of the Sentiment Analysis Techniques

Geetika Vashisht, Manisha Jallia

Abstract— Sentiment analysis (SA) is a rapidly evolving field that aims at computationally categorizing the opinions of people about a particular product, movie, brand or anything that can be opined. It has changed the way the information is perceived and utilized by big business groups, brands and marketing agencies by demonstrating that the computational recognition of a sentimental expression is feasible. The fruition of Internet based applications has generated huge amount of personalized views on the Web. These reviews exist in different forms like social Medias, blogs, Wiki or forum websites. The boom of search engines like Yahoo and Google has flooded users with copious amount of relevant reviews about specific destinations, which is still beyond human comprehension. Sentiment Analysis poses as a powerful tool for users to extract the needful information, as well as to aggregate the collective sentiments of the reviews. This research will explore and compare the various techniques used for Sentiment Analysis in the last decade.

Keywords— Sentiment Analysis(SA); Sentiments; Lexicons; Machine Learning(ML)

I. INTRODUCTION

"Emotions can get in the way or get you on the way."
-Mavis Mazhura

This famous quote by Mavis highlights the importance of Emotions in our lives! Emotions a.k.a. sentiments are expressed in myriad ways today and unlike the traditional era, where expressing one's opinion was restricted to verbal communication; today Internet has revolutionized how we express ourselves. Social networking sites like Facebook, twitter, Instagram; review sites like mouthshut, glassdoor; blogs and several similar platforms are easily accessible for voicing one's opinion. These sentiments thus expressed work as gold mines for businesses, marketing agencies, stock traders. These days even the political parties are actively using the E-Word Of Mouth (e-wom) to identify the grey areas and to even define their propaganda. The world has come closer with each individual contributing to the overall sentiment score. Isn't that great to participate and collaborate on such a grand scale. This led to the inception of a very intriguing field of sentiment analysis a.k.a. opinion mining. Sentiment analysis is a technique of turning unstructured information into structured data that classifies attributes of the text viz. sentiment orientation (positive or negative or neutral) of opinions, the object that is being discussed and the entity that expresses the sentiments. The knowledge of linguistics contributes greatly in mining the sentiments and is inevitably a pre-requisite to work with lingual text.

Sentiment Analysis is so tightly coupled with the field of statistics that statisticians sometimes refer to it as "applied statistics". Statistical methods are required to do the exploratory data analysis and to derive meaningful correlation between data items and to turn raw observations into information. Domain Expertise is indeed another vital aspect that can't be ignored. Unless and until we are not well acquainted with the domain under consideration, it's not possible to filter the exact data that will contribute to consequential results. The market around SA has matured up to the point that there are many products out there to do the analysis of sentiments. Approximately 2.5 quintillion bytes of data is generated everyday and is multiplying at a faster pace cropping up the need to develop tools and techniques to make sense of that data. Different levels at which sentiment analysis is performed is discussed in section II followed by the discussion on the techniques for mining sentiments in section III. Section IV compares the performances of different approaches and presents the classified work of reported literature in the field. Section V highlights the important existing tools. Section VI identifies the major challenges faced by the field of sentiment analysis followed by conclusion in section VII.

II. LEVELS OF SENTIMENT ANALYSIS

Sentiment analysis has been investigated mostly at five levels. (Ge, Vazquez, & Gretzel, 2018)

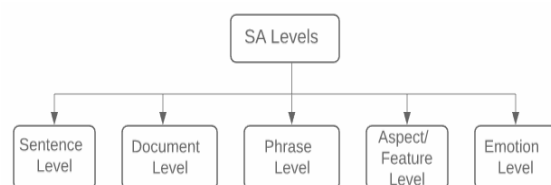


Fig. 1 : Levels of Sentiment Analysis

Sentence Level

Sentence level analysis aims at identifying the polarity of a sentence thus is widely applied to microblogging posts. This kind of analysis segregates objective sentences having concrete information from the subjective ones having subjective views and opinion. (Devika M D, 2016) According to (Kolkur, Gayatri, & Mahe, 2015), sentence level analysis comprises of two tasks - subjectivity classification and sentiment classification.

Document Level

The whole "emotional" document is under scrutiny in Document-level analysis and the overall polarity (positive or negative) of the document is identified (Bo Pang and Lillian Lee, July 2002).

Panoptical View of the Sentiment Analysis Techniques

This kind of analysis has proven to be extremely beneficial for recommender systems to get an insight of the mass opinion (B.Pang, 2004). Online reviews, blogs, articles, editorial sites, forum discussions are the key sources for applying document level analysis. Though there's a limitation on performing the fine-grained tasks but it's simplest way to classify the overall polarity of a text. (Liu B.,2015). Document-level analysis is not applicable to documents containing opinions about multiple entities as it strictly focuses on a single entity posing a serious limitation in the real world analysis as seldom a document focuses on a single topic(Yi, Nasukawa, Bunescu, & Niblack, 2003). When there is a need to compare several topics discussed in a document, this analysis fails badly as in-depth analysis is not possible.

Phrase Level

Phrase-level analysis targets contextual polarity of a text. A word having a positive meaning but the overall contextual polarity of the phrase in which the word appears can vary. For example, "Without the centralised heating systems, this place is too cool". Now, "cool" generally indicates positive sentiment but in the given context, the polarity of the phrase is not positive. Thus, the polarity of the phrase is not just dependent on the opinion words but on the word's contextual polarity. (Wilson, Wiebe, & Hoffmann, 2009) devised a mechanism to automatically identify the contextual polarity for a sentiment expressions by first determining whether an expression is polar or neutral and disambiguating the polar ones.

Aspect Level

Also known as feature-based, topic-based, entity-based and target-based analysis; the focus here shifts from language constructs (documents, paragraphs, sentences, clauses or phrases) to identifying aspects of an entity and associated sentiment words. The analysis starts with identification of the entity under observation. Sentiment words related to the entity are identified and categorised as per the polarity (positive, neagtive and neutral). Thus, the aspects related to the entity are vividly discovered and scored as per polarity, achieving fine-grained analysis (Mohan, R, B, & J, 2014) For example, consider the review about a hotel- "The hotel room is very clean.". Here, room is the aspect (noun) and 'clean' is the opinion word (adjective) (Tribhuvan, Bhirud, & Tribhuvan, 2014)

Emotion Level

Emotion level categorises the text on the basis of the underlying emotions like happy, sad, anger or joy. This kind of analysis is more detailed and accurate. Emoticons(Emojis) based sentiment analysis works at this level. It's easier to classify the emoticons(emojis) and thus the emotions or the sentiments. Emotion level analysis is often reliant on the Lexicon based approach. (Ge, Vazquez, & Gretzel, 2018).

Further, the methods to measure the sentiment strength can be categorized based on the rating levels - one for identifying aspects of a product or service and the other to rate a review on a global level that considers only the polarity of the review (positive/negative)(Anaisi Collomb).

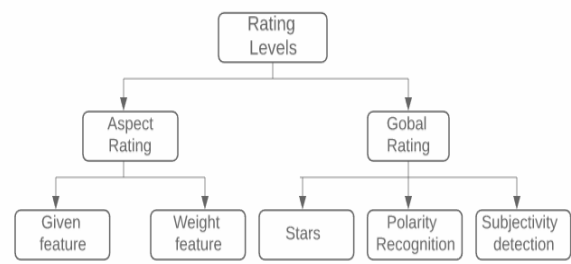


Fig. 2 : SA Rating Levels

III. SENTIMENT ANALYSIS APPROACHES

There are many methods to carry out sentiment analysis (fig. 3). Research is still going on to find out better alternatives due to the importance of sentiment analysis in today's global economy..

A. Machine Learning Approach

Machine learning strategies have speed up the once gruelling data crunching tasks. It works by training an algorithm with a training data set developed from examples and experiences. This means the algorithm learns automatically from the data that is provided and this is called the training phase of the algorithm followed by a testing phase where we feed some test data to the algorithm to get the expected outcome. The approach can be further categorised as Supervised and unsupervised.

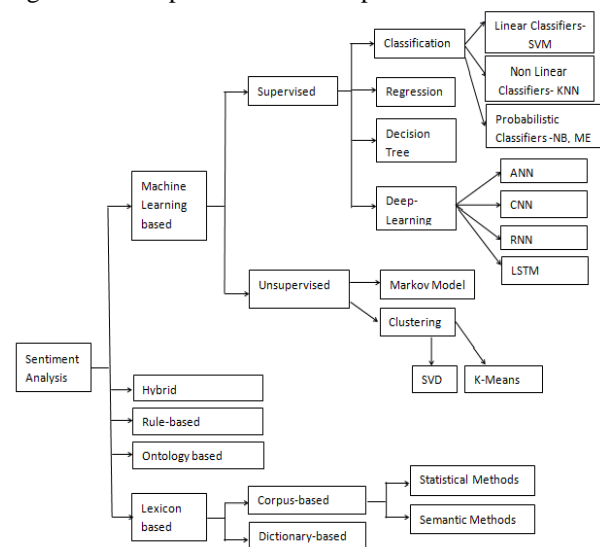


Fig. 3 : Sentiment Analysis Approaches

Supervised Machine Learning Approach

In this approach, given a sample of training data, the algorithm continuously predicts the outcome till the acceptable level of performance is achieved. There is prior knowledge of the outcome. Supervised learning is commonly done in context of classification, when we want to map categorical input data into labelled classes or regression, where input data is mapped to continuous output rather than discrete. The underlying motive is to find the specific co-relation in the input dataset to predict the correct outcome.



Supervised Learning	
Classification	Regression
Fraud Detection	Weather forecasting
Image Classification	Population growth, Salary & real estate prediction
Customer Sentiment Analysis	Financial Portfolio Prediction
Diagnostics	Market Trends

Fig. 4 : Applications of Supervised Learning

1.1) Classification

Classification is the most common task in sentiment analysis. It's a supervised machine learning technique for classifying the polarity of the emotional data. There are several popular approaches of classification. Support Vector Machine (SVM), a supervised linear classifier, can be implemented for both classification and regression. Works really well in text classification problems with extremely higher dimensions. KNN is most widely used classification model and belongs to the supervised learning domain but can be used for regression too. It classifies by seeking the similarity with the ones already classified and tagged. A new object is assigned to the class which has most common neighbours among its 'k' nearest neighbours where 'k' is a typically small positive integer. Naive Bayes is a probabilistic classifier for developing classifiers with a basic assumption of independence between predictors. Naive Bayes computes class probabilities and conditional probabilities from the dataset. It uses the bayes theorem that finds the probability that an event will occur given the probability that another event has already occurred. The Max Entropy classifier is again a probabilistic classifier that is not based on the assumption that the features are conditionally independent of each other. It is based on the Principle of Maximum Entropy and from all the models that fit the training data, it chooses the one having the largest entropy. Unlike Naive Bayes, that allows each feature to have its say independently, Max Entropy classifier uses search-based optimization to get weights for the features that maximize the probability of the training data. This When there is little knowledge about the prior distributions, it's challenging to implement Maximum Entropy classifier as it works on minimum assumptions and it is not practical to make any random assumptions.

1.2) Regression

Regression works on continuous data or we can say real numbers whereas if the data is unordered countable and distinguishable then classification is used.

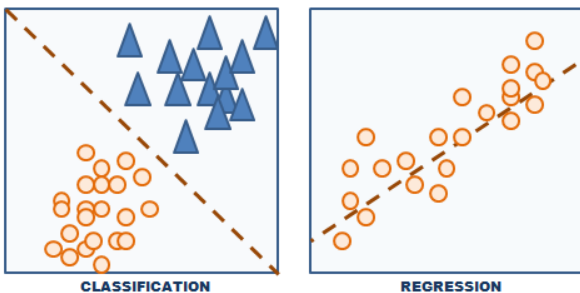


Fig. 5 : Classification Vs. Regression

1.3) Decision Tree

Like other classification models, decision trees tries to divide data into two or more sets. The idea is to find out the most significant variable that can create the best split. Consider the example to distinguish between Two persons A and B using the variables weight, height and dress colour. From a set of images, decision tree would split the data on each of the variable to find the best one out. Both persons don't have much difference in height and may wear same colour shirts too but there is considerable difference in their weights leading the algorithm to choose weight as the split variable (fig 1.)

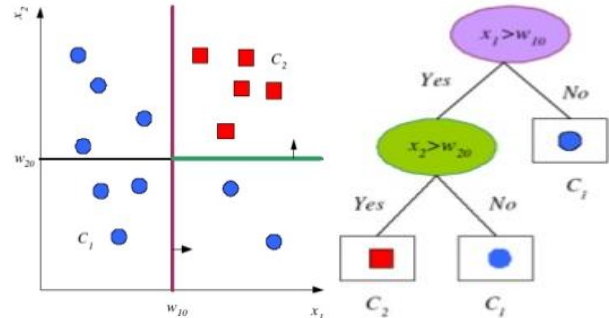


Fig. 6 : Illustration of Decision Trees - Nearest Neighbours

1.4) Deep Learning

Deep Learning approaches have surfaced as powerful computational models and have brought a revolution in the field of Sentiment Analysis. Deep Learning exploits the power of neural networks which consists of a large number of neurons(the information processing units) organised in layers. The simple structure of a neural network is shown in the figure below:

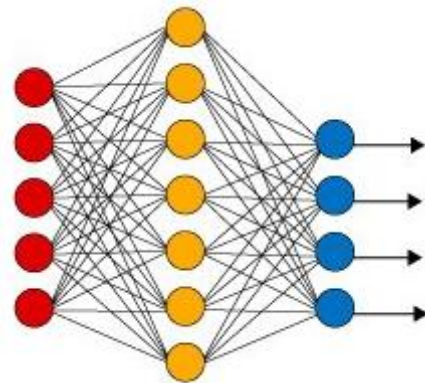


Fig. 7 (a) : A three layer simple neural network-one hidden layer of seven neurons(units) and one output layer with four neurons and five inputs

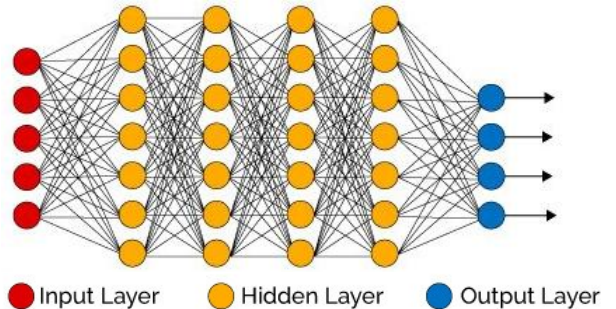


Fig. 7 (b) : Deep Learning neural network (Kampakis, 2018)

Panoptical View of the Sentiment Analysis Techniques

A three layer neural network with five inputs, four hidden layers of seven neurons(units) each and one output layer. Notice that in both the cases, there are connections(synapses) between the neurons across the layers, but not within a layer.

Unsupervised Machine Learning Approach

When dealing with real world problems, the associated data will not come with pre-defined labels which is a pre-requisite in Supervised machine learning algorithms. In this case, unsupervised machine learning models come into picture by classifying data on the basis of some commonality in the features which is then used to predict the classes on new data. Aim is to identify the natural structure or distribution in the data present within a set of data points without using explicitly provided labels. In unsupervised learning methods, it's challenging to compare the model performance as no labels are provided. Still, it finds its applications in dimensionality reduction and exploratory analysis. The overall process is summarised in the figure .

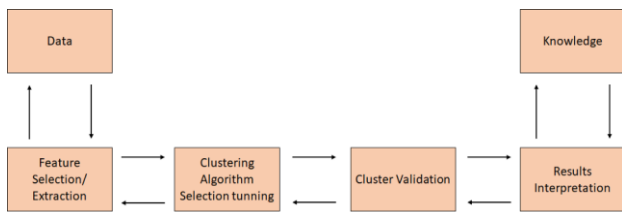


Fig. 8 : Unsupervised Learning Analysis Process

Apart from segregating datasets, unsupervised machine learning helps identify the anomalies or the outliers. Other common applications are presented in figure .

Unsupervised Learning

	Clustering	Dimensionality Reduction
Applications	Recommender Systems	Big data visualization
	Outliers detection- Spam filters	Meaningful compression
	Marketing & sales	Structure Discovery
	Classifying network traffic	Neuroscience- Maximally Informative dimensions
	Identifying fake news	Feature Elicitation

Fig. 9 : Applications of Unsupervised Learning

1.1) Hidden Markov Model

This probabilistic model predicts a sequence of hidden variables from a set of observed variables. It finds its applications in detecting facial expressions from videos and hence the sentiment behind it.

1.2) Clustering

Clustering algorithms are preferred when we want to divide the data on the basis of some common attributes. If there is a problem of segregating something or concentrating on particular groups then Clustering is chosen. The simplest algorithm to solve clustering problems by partitioning the n observations into k clusters. The major challenge is that K-means needs to know the number of clusters in the data in advance so it can be tricky to get the best value for "k".

Another approach in the league is SVD (Singular Value Decomposition). It is based on a theorem in linear algebra and is used for aspect based sentiment analysis. SVD provides numerous ways of diagonalizing a matrix into special matrices that are straightforward to govern and to analyse. It conjointly lay down the basis to transform the information into independent components. As SVD can find the latent relation among the data, (Thara.S, 2017) used SVD based feature for sentiment prediction. Limited research has been conducted on application of SVD in sentiment analysis.

B. Hybrid Approach

Some researchers have concluded that the combination of both the lexicon based approach and machine learning approaches gives more accurate results. A hybrid approach combines machine learning and the lexicon-based approaches for sentiment analysis. It uses the lexicon-based approach for calculating the sentiment score followed by training a classifier. It achieves high accuracy provided by a powerful supervised learning algorithm and stability by lexicon based approach.

C. Rules based Approach

The rule based sentiment analysis approaches are supervised techniques comprising of three phases, the first being the product feature extraction followed by opinion sentence extraction and opinion orientation identification.

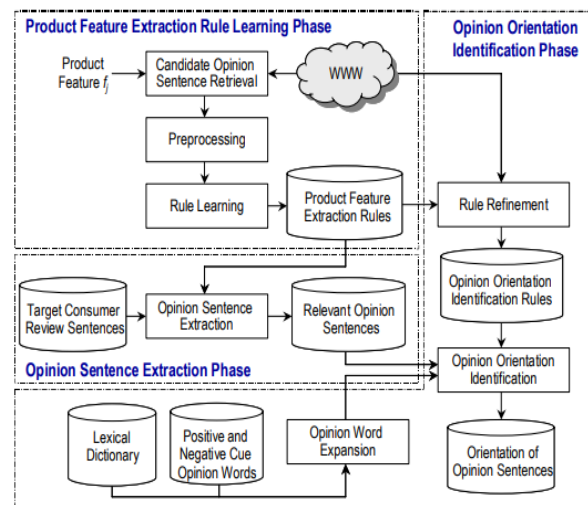


Fig. 10 : Illustration of Rule based Sentiment Analysis Technique (Yang & Shih, 2012)

Like supervised learning approaches, there is no need to prepare training data manually. A practical approach to solve real world problems with numerous entities having copious features.

D. Ontology Based Approach

Ontology based sentiment analysis works by exploring the domain knowledge and identifying the relationship that exists between the objects in the domain. Many researchers have worked on identifying the features of a product to develop an ontology model. (Haider, 2012) and (Yaakub & Feng, 2011) came up with an ontology model to analyse customer's sentiments on mobile phones based on the products' features. (Tim Finin



& Zou, 2005) came up with an innovative mechanism of developing an ontology based intelligent application.

E. Lexicon Based Approach

These approaches are based on external lexical resources that map words in the lexicon to a categorical (neutral, negative, positive) or numerical sentiment score. This score is used by the algorithms to obtain the overall sentiment conveyed by the text. The efficacy of the approach is based on the quality of the lexical resource. It doesn't require any training data like the Machine Learning approach but to come up with an exhaustive lexicon is always a challenge due to the versatile and ever changing nature of the expressions and natural language. Most common lexical resources are SentiWordNet, WordNet-Affect, MPQA and SenticNet (concept-level SA). Lexicon based approach can be further divided into Dictionary-based and Corpus-based.

In Dictionary-based approach, firstly the opinion word from review text are found, which is followed by finding their synonyms and antonyms from dictionary. The dictionaries like WordNet, SentiWordNet, SenticNet may be incorporated for mapping and scoring.

Corpus-based method helps to find opinion word in a context specific orientation. Beginning with a list of opinion word, the corpus-based approach finds other opinion word in a huge corpus..

IV. COMPARISON AND CONSOLIDATION

Table 1 highlights the criteria to decide the approach that can be opted for a particular problem at hand. As per the research, Machine Learning Approach is the most promising when it comes to accuracy while the speed of execution is achieved by the rule based approach.

TABLE I WHEN TO USE WHAT!

Criteria	Rules	Machine Learning	Lexicons
Accuracy		✓	
Maintenance		✓	✓
Interpretability	✓	✓	✓
Speed of execution	✓		

TABLE III

UPSUM OF THE APPROACHES AND EVALUATION METRICS USED BY AUTHORS

Author(s)	Dataset	Technique (Evaluation)
(Read, 2005)	Newswire	Accuracy: Topic dependency: NB Accuracy - 75.5 to 84.6% SVM Accuracy - 75.5 to 81.1% Domain dependency: NB Accuracy - 78.2 to 78.9% SVM Accuracy - 78.2 to 81.5% Temporal dependency:

		NB Accuracy - 76.5 to 78.9% SVM Accuracy - 80.8 to 81.5%
(K. Dave, 2003)	Amazon, CNET	Accuracy: NB :81.9%-87% SVM:85.8% - 87.2%
(Alec Go, 2009)	1,600,000 Tweets	Accuracy: NB: 82.7% Max-Ent: 82.7% SVM: 82.2%
(Dmitry Davidov, 2010)	475 million Tweets	Precision : 90 - 92%
(L. Zhang R. G., 2011)	Tweets	Hybrid Approach Accuracy: 85.4%
(Liu M. H., 2004)	Amazon, CNET	Accuracy : Lexicon: 84%
(Agarwal, 2011)	11,875 Tweets	Accuracy : 60.83%
(A. Abbasi, 2008)	U.S. & Middle Eastern web forum postings	Accuracy: SVM : 95.55%
(A. Mudinas, 2012)	CNET, IMDB	Hybrid Approach Accuracy: 82.30%
(Kouloumpis, 2011)	22,2474 tweets	Avg. Accuracy: HASH : 74% HASH+EMOT:75% Avg. F-measure: HASH : 68% HASH+EMOT:65%
(Chen, 2011)	Multi-Domain	Hybrid Approach Accuracy: 66.8%
(Yuita Arum Sari, 2014)	WordNet-Affect, WPARD datasets (Indonesian language)	Accuracy: KNN : 92%
(Bo Pang and Lillian Lee, July 2002)	IMDB	Accuracy: NB : 81.5%
(Turney, 2002)	Epinions	Accuracy: PMI : 66%
(A. Khan, 2011)	IMDB, Skytrax, Tripadvisor	Accuracy : Lexicon: 86.6%



Panoptical View of the Sentiment Analysis Techniques

TABLE III
CLASSIFIED WORK OF REPORTED LITERATURE IN THE FIELD OF
SENTIMENT ANALYSIS

Author(s)	SA Level	M L	Rules based	Lexi-con Based	Hyb -rid
(VOHRA & TERAIYA, 2013)	✓	✓		✓	✓
(Medhat, Hassan, & Korashy, 2014)	✓	✓	✓	✓	✓
(Kolkur, Dantal, & Mahe, 2015)	✓	✓			
(D'Andrea, Ferri, Gifroni, & Guzzo, 2015)	✓	✓		✓	✓
(Schouten & Frasincar, 2015)	✓*	✓	✓	✓	✓
(Bhadane, Dalal, & Doshi, 2015)		✓		✓	
(Khan, Durrani, Ali, Inayat, Khalid, & Khan, 2016)	✓	✓		✓	
(Devika M D, 2016)	✓	✓	✓	✓	✓
(Ahlgren, 2016)	✓	✓			
(Rajput & Solanki, 2016)		✓		✓	
(Purvi Prajapati, April 2017)	✓	✓	✓	✓	
(Kathuria & Upadhyay, 2017)	✓	✓	✓	✓	
(Kakulapati, 2017)	✓	✓		✓	
(Zia, Fatima, IdrisMala, Khan, Naseem, & Das, 2018)	✓	✓		✓	
(Wankhede Rohit, 2018)	✓	✓			
(Redhu, Srivastava, Bansal, & Gupta, 2018)	✓	✓	✓	✓	✓
(Pandi, Dandibhotla, & Bulusu, 2018)	✓	✓		✓	
(Ge, Vazquez, & Gretzel, 2018)	✓	✓		✓	
(Anais Collomb)	✓*	✓	✓	✓	
(Mäntylä, Graziotin, & Kuutila, 2018)	✓	✓		✓	

✓* includes rating levels also.

V. POPULAR SENTIMENT ANALYSIS TOOLS

Some of the common tools in the repertoire of social media analysis done by various business groups are Brand24,

MeltWater, Google Alerts, Google Analytics, Tweetstats, Facebook Insights, Social Mention, Clarabridge, Repustate, OpenText, Lexalytics and the list goes on!

TABLE IVV
SENTIMENT ANALYSIS TOOLS (ARVINDER KAUR, 2016) (ALESSIA D'ANDREA, 2015)

Tools	Technique Used
Senti WordNet	Hybrid approach(Lexicon dictionary and machine learning based techniques)
SenticNet	NLP
EMOTICONS	Emoticons based SA
Basis Technology Rosette	Hybrid approach(Lexicon dictionary and machine learning based techniques)
Text Sentiment Visualizer	Deep Learning
NRC	Lexicon based
FRN	n-gram
Dataladder productmatch	Machine Learning
Lexalytics	Rules based and Machine Learning
MonkeyLearn	Machine Learning
Megaputer Text Analyst	Machine Learning
Netowl	Machine Learning
SIFT	Machine Learning
Meaningcloud	Feature-level SA
Data Science Toolkit	Machine Learning

VI. MAJOR CHALLENGES IN SENTIMENT ANALYSIS

Subjective and opinionated data is multiplying at lightning speed with an equal increase in the distinctive potential data sources. This has brought about critical challenges to the field of sentiment analysis:

Multimodal communication : E-users can use text, videos, emoticons(emojis), pictures or any other graphic icons to express themselves online. This increase in the dimensions of the opinionated data complicates the process. Popular social networking sites like Facebook provides users with myriad of emoticons manifesting kinetic gestures and graphic icons to express themselves. These emoticons are culture-specific and keeps on changing with time posing a challenge to interpret the underlying sentiments.

Multilingual text : A text can be multilingual. For example, consider the review- "Mast picture thi. Loved it. <3". The review is in two languages viz. hindi and english and the number of languages can be even more posing a challenge to mine the underlying sentiments.

Intertextual text : Intertextual text is hard to interpret as it is interconnected and dependent on some other text in the form of hyperlinks, reposts or embedded images. thus, the meaning of the text is hidden in several layers of texts. Analysing sentiments in this case can be quite tricky.

Managing sarcasm : Natural Language can be hard to interpret by an algorithm especially when sarcasm and wits have been used. For example, consider the review- "Thanks to the weather that I couldn't step out of the house.". Humans can easily sense sarcasm but still sentiment analysis approaches need to train themselves here.

Domain, Topic and Temporal dependency : Machine Learning models are domain-dependent, topic-dependent and temporally dependent. For example, consider the review - "Kashmir is beautiful but scary too." The reviews can be biased after terrorist activities in a particular region. If a model is trained on the basis of reviews in that particular time span then the model will never approve of Kashmir as a good tourist place. The amount of time over which a model shall be trained to avoid temporal dependence is still an open issue. (Read, 2005) carried out experiments to determine the influence of domain, topic and time on machine learning based sentiment classification. Training a classifier using emoticons together with text could function independent of domain, topic and time. Emoticons are a great way of overcoming the domain, topic and time problems.

Word Sense Disambiguation (WSD) : WSD is a technique to discover the ambiguity which crops up due to multiple meanings of words in a different context. Take for instance, the two sentences. "The *bank* will not be accepting cash on Saturdays." and "The river overflowed the *bank*." In the former sentence, the word *bank* refers to the commercial money banks and in the latter the river's bank is mentioned. This is challenging for an algorithm to detect this.

Handling spam opinions : Marketers flood fake opinions to give a boost to their products. Identifying spam or fake opinions is still a challenge to find a detection algorithm.

VII. CONCLUSION

This paper examines the sentiment analysis process, its different levels of analyzing sentiments viz. Sentence level, Document Level, Phrase Level, Aspect Level, Emotion Level followed by approaches to analyze sentiments viz. Machine Learning Approaches, Lexicon based approaches, Hybrid approaches, Rules based approaches, ontology based approaches. An effort has been made to cover all the approaches in the field till date. Classified work of the reported literature in the field has been presented with an aim to cover all the best cited papers concluding that the success of an approach depends majorly on the problem at hand. Additionally, an amalgam of rule based approach, Machine learning approach and lexicon based approach can be promising option. Sentiment analysis tools and the underlying technique they incorporate is covered. Though a lot of research has been done in the field of Sentiment Analysis, it is yet to mature. Several challenges and open issues are discussed at length. In this E-era, Sentiment Analysis has spread its wings in every one's life, directly or indirectly we all are contributing to this extremely fascinating field of mining sentiments.

REFERENCES

1. Anais Collomb, A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation.
2. Antonio Moreno-Ortiz, J. F.-C. (2015). Identifying polarity in financial texts for sentiment analysis: a corpus-based approach. *Procedia - Social and Behavioral Sciences* 198 , 330-338.
3. Chetashri Bhadane, H. D. (2015). Sentiment analysis: Measuring opinions. Elsevier *Procedia Computer Science* 45 , 808-814.
4. Deepa Ananda, D. N. (2016). Semi-supervised Aspect Based Sentiment Analysis for Movies using Review Flitering. 7th International conference on Intelligent Human Computer Interaction, IHCI 2015, *Procedia Computer Science* 84 , 86 – 93.

5. Deepali Arora, K. F. (2015). Consumers' sentiment analysis of popular phonebrands and operating system preference using Twitter data: A feasibility study. *IEEE 29th International Conference on Advanced Information Networking and Applications* , 680-686.
6. Devi, M. U. (2013). Analysis of sentiments using unsupervised learning techniques. *Information Communication and Embedded Systems (ICICES)* , 241-245.
7. Devika M D, S. C. (2016). Sentiment Analysis:A Comparative Study On Different Approaches. *Procedia Computer Science* 87 , 44 – 49.
8. Feldman, R. (2013). Techniques and Applications for Sentiment Analysis. *Communications of the ACM* , 82-89.
9. Funk, D. M. (2012). Automatic detection of political opinions in tweets. *The Semantic Web: ESWC 2011 Workshops*, ser. Lecture Notes in Computer Science, R. G. Castro, D. Fensel, and G. Antoniou Eds., vol. 7117. Springer Berlin Heidelberg, , 88-89.
10. L. Zhang, R. G. (2011). Combining lexicon-based and learning-based methods for twitter sentiment analysis. [Online]. <http://www.hpl.hp.com/techreports/2011/> .
11. Li, H.-X. S.-J. (2011). sentiment analysis model for hotel reviews based on supervised learning. *Machine Learning and Cybernetics (ICMLC)*, 2011 International Conference on, vol. 3 , 950-954.
12. Liu, G. L. (2010). A clustering-based approach on sentiment analysis. *Intelligent Systems and Knowledge Engineering (ISKE)*, 2010 International Conference , 331-337.
13. Liu, M. H. (2004). Mining and summarizing customer reviews. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '04.NY, USA: ACM , 168-177.
14. Matteo Cristani, C. T. (2018). Making sentiment analysis algorithms scalable. *Research Gate*.
15. Sharma, S. (2018). A Context Based Algorithm for Sentiment Analysis. *Research Gate Article in International Journal of Computational Vision and Robotics* .
16. Siau, E. A. (2014). An approach to sentiment analysis the case of airline quality rating. *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)* , 363–368.
17. Sowmya Kamath S, A. B. (2013). Sentiment Analysis Based Approaches for Understanding User Context in Web Content. *International Conference on Communication Systems and Network Technologies* , 607-611.
18. Theodoros Giannakopoulos, M. P. (2015). Visual sentiment analysis for brand monitoring enhancement. *9th International Symposium on Image and Signal Processing and Analysis (ISPA 2015)* .
19. Wala Medhat, A. H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal* , 1093-1113.
20. Wang, W. (2010). Sentiment analysis of online product reviews with semisupervised topic sentiment mixture mode. *Fuzzy Systems and Knowledge Discovery (FSKD)*, 2010 Seventh International Conference vol 5 , 2385-2389.
21. Wankhede Rohit, R. J. (2018). An Approach to Sentiment Analysis. *JSRST* , 1508-1513.
22. Y. Singh, P. K. (2007). A review of studies on Machine Learning Techniques. *International Journal of Computer Science and Security* , 70-84.
23. Yenier Castañeda, A. C. (2014). UMCC_DLSI_Prob: A Probabilistic Automata for Aspect Based Sentiment Analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)* , 722–726.
24. BIBLIOGRAPHY (n.d.). Retrieved April 2019, from English Grammar Today: <https://dictionary.cambridge.org/grammar/english-grammar-today/>
25. Alexander Hogenboom, D. B. (2013). Exploiting Emoticonsin Sentiment Analysis. *ACM* , 18-22.
26. B. Jansen, M. Z. (2009). Twitter Power: Tweets as Electronic Word of MOUTH. *Journal of the American Society for Information Science and Technology* , 2169-2188.
27. Bickart, B. &. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing* , 31-40.
28. Dmitry Davidov, O. T. (2010). Enhanced Sentiment Learning Using Twitter Hashtags and Smileys. *Proceedings of the 23rd international conference on computational linguistics: posters*. Association for Computational Linguistics. , 241-249.
29. Fahlman, S. (1982, September 19). Retrieved from <http://www.cs.cmu.edu/~sef/Orig-Smilely.htm>
30. Gaël Guibon, M. O. (2016). From Emojis to Sentiment Analysis. *WACAI 2016* , <<http://www.enib.fr/wacai/>>.

31. Lohr, S. A \$1 Million Research Bargain for Netflix and Maybe a Model for Others. The New York Times.
32. P. Kralj Novak, J. S. (2015). Sentiment of emojis. . PLoS ONE 10(12): e0144296. .
33. Read, J. (2005). Using Emoticons to reduce Dependency in Machine Learning Techniques for Sentiment Classification. Proceedings of the ACL Student Research Workshop, ACM , 43-48.
34. Richard D. Lawrence, V. S. (2009). Social Media Analytics: Channeling the Power of the Blogosphere for Marketing Insight. Retrieved from Research Gate: https://www.researchgate.net/publication/255623032_Social_Media_Analytics_Channeling_the_Power_of_the_Blogosphere_for_Marketing_Insight/
35. Rothenberg, M. (n.d.). Retrieved April 2019, from emojitracker: <http://emojitracker.com/>
36. Shivani Bahria, P. B. (2018). A Novel approach of Sentiment Classification using Emoticons. International Conference on Computational Intelligence and Data Science , 669-678.
37. Wright, A. (2009). Mining the Web for feelings Not facts. New York Times.
38. Yuita Arum Sari, E. K. (2014). USER EMOTION IDENTIFICATION IN TWITTER USING SPECIFIC FEATURES:HASHTAG, EMOJI, EMOTICON, AND ADJECTIVE TERM. Journal of Computer Science and Information , 18-23.
39. BIBLIOGRAPHY \1 1033 Alexander Hogenboom, D. B. (2013). Exploiting Emoticons in Sentiment Analysis. ACM , 18-22.
40. B. Jansen, M. Z. (2009). Twitter Power: Tweets as Electronic Word of MOUTH. Journal of the American Society for Information Science and Technology , 2169-2188.
41. Fahlman, S. (1982, September 19). Retrieved from <http://www.cs.cmu.edu/~sef/Orig-Smile.htm>
42. Gaël Guibon, M. O. (2016). From Emojis to Sentiment Analysis. WACAI 2016 , <<http://www.enib.fr/wacai/>>. <hal-01529708>.
43. Read, J. (2005). Using Emoticons to reduce Dependency in Machine Learning Techniques for Sentiment Classification. Proceedings of the ACL Student Research Workshop, ACM , 43-48.
44. Richard D. Lawrence, V. S. (2009). Social Media Analytics: Channeling the Power of the Blogosphere for Marketing Insight. Retrieved from Research Gate: https://www.researchgate.net/publication/255623032_Social_Media_Analytics_Channeling_the_Power_of_the_Blogosphere_for_Marketing_Insight/
45. Rothenberg, M. (n.d.). Retrieved April 2019, from emojitracker: <http://emojitracker.com/>
46. Shivani Bahria, P. B. (2018). A Novel approach of Sentiment Classification using Emoticons. International Conference on Computational Intelligence and Data Science , 669-678.

AUTHORS PROFILE



Geetika Vashisht received the Masters in Computer Applications degree from Guru Gobind Singh Indraprastha University in 2010 and BSc.(H) Computer Science degree from the University of Delhi in 2007. In 2010, she joined IT industry and worked with big brands such as Samsung till 2013. Later, she cleared UGC_NET and joined as an Assistant Professor in Delhi

University in 2013. Her current research interests include Data Mining, Sentiment Analysis, Data Science. She is a lifetime member of Computer Science Society of India, Delhi Chapter.



Manisha Jaillia received doctorate of Philosophy from Banasthali Vidyapith. and is currently working as a Sr. Assistant Professor (CS) Banasthali University, Tonk (Raj.) Her research area includes data mining and software engineering.